

09/720934

**SH3D1A Domain Structure and Homologies - Human vs Xenopus
(Determined using GCG programs, BLAST, FASTA)**

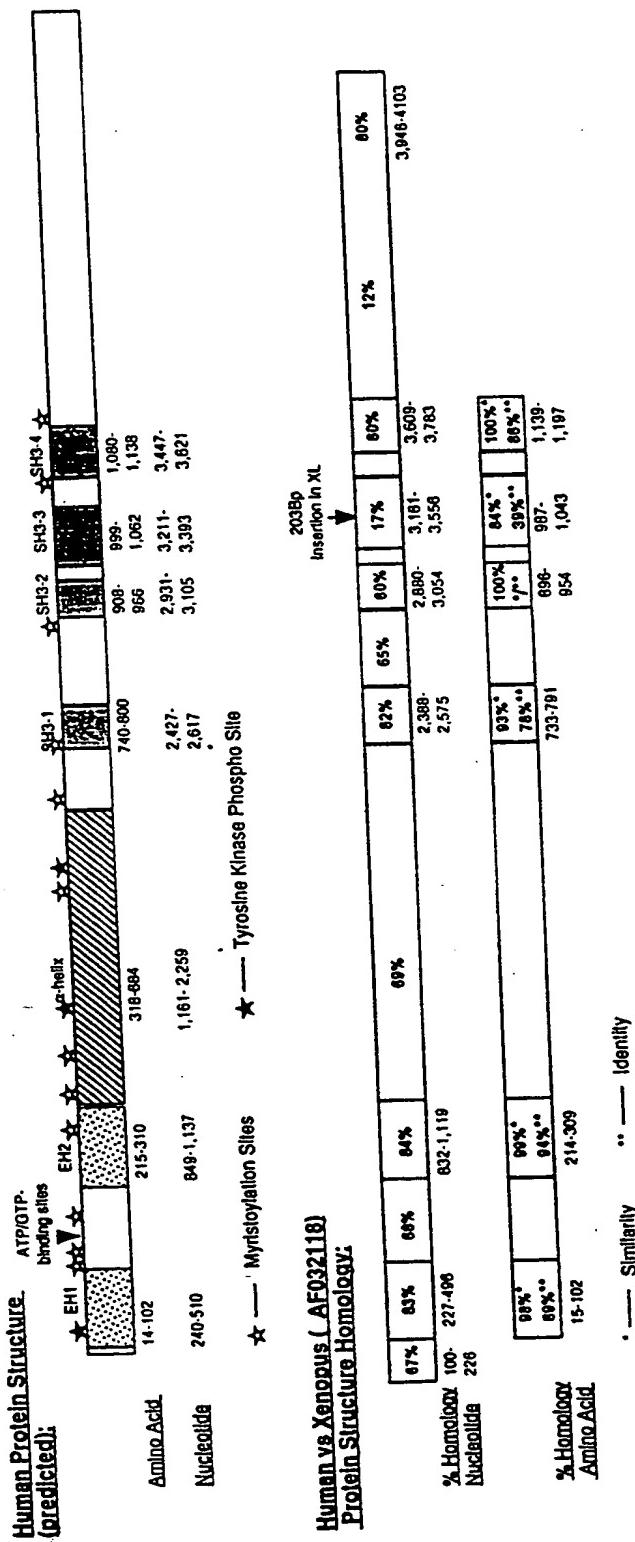
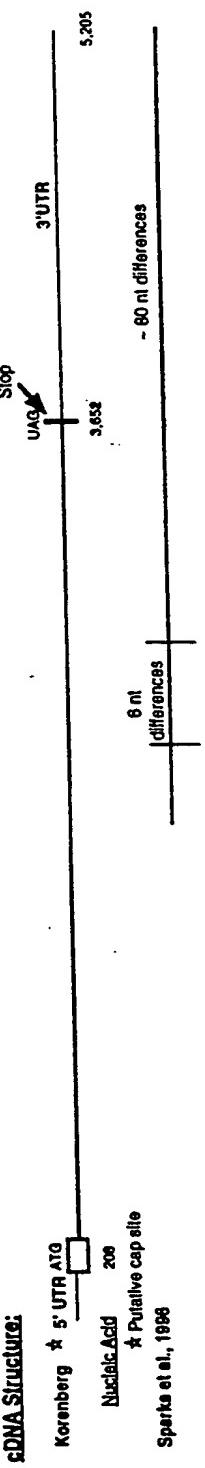
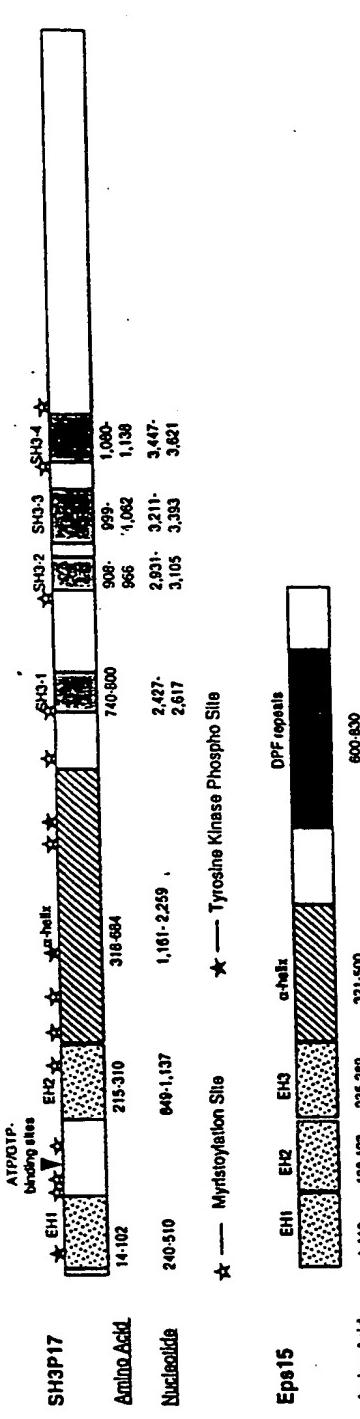


Figure 1

Human SH3D1A Structure and Homology

**EH Domain Comparison of Human SH3PP17 and Mouse Eps15,
(Epidermal growth factor (EGF) receptor pathway substrate clone 15):**



cale for Above Ideograms:

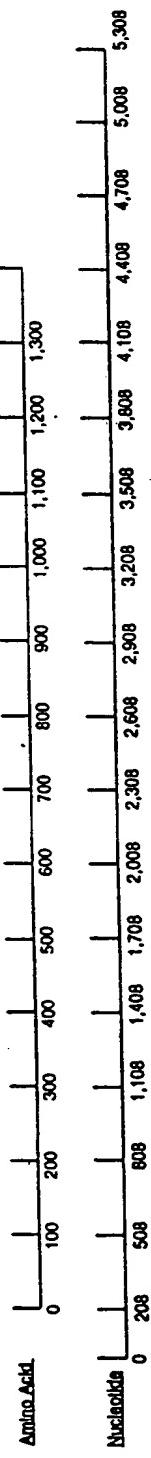


Figure 2

Region of Chromosome 21 Responsible for Megakaryocytic Abnormalities

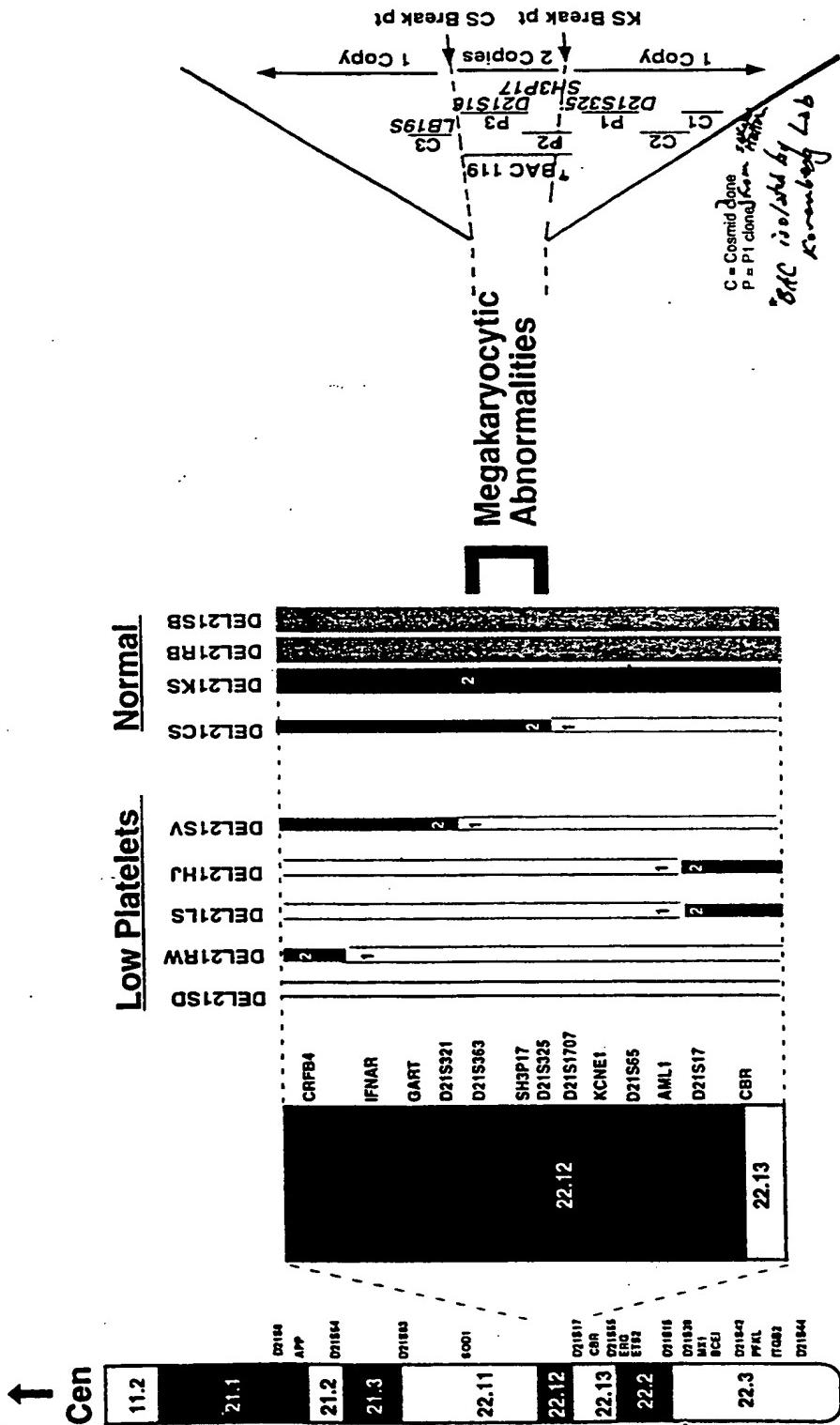


Figure 3

09/720934

WO 99/53062

PCT/US99/08371

4 / 30

SH3D1A

1 CAAAAGAAATT CGGGTACGG CGGCTCGOGA GGAAGAAATCC CGAGCGGGCT
 51 CGGGGACGGA CAGAGACCGG GGCGGGGATG GTGTGCGGGG CTGCGGCCTC
 101 TGGGTCCCTC CCAGCGGCGC GTGACGGGCA CTGATTGTC CCTGGGGGOGG
 151 CAGCGOOGGAC CGGCGGGAG ATGAGGGGTC GATTAGCAAG GTAAAAGTAA
 201 CAGAACCATG GCTCAGTTTC CAACACCTTT TGGTGGCAGC CTGGATAATCT
 251 GGGCCATAAC TGTAGAGGAA AGACCGAACG ATGATCAGCA GTTCCATAGT
 301 TTAAAGCCAA TATCTGGATT CATTACTGGT GATCAAGCTA GAAACTTTTT
 351 TTTTCAATCT GGGTACCTC AACCTGTTT ACCACAGATA TGGGCACTAG
 401 CTGACATGAA TAATGATGGA AGAATGGATC AAGTGGAGIT TTCCATAGCT
 451 ATGAAACTTA TCAAACATGAA GCTACAAGGA TATCAGCTAC CCTCTGCACI
 501 TCCCCCTGTC ATGAAACAGC AACCAGTTGC TATTTCTAGC GCACCAGCAT
 551 TTGGTATGGG AGGIAATGCC AGCATGCCAC CGCTTACAGC TGTGCGCTCA
 601 GTGCCAATGG GATOCATTC AGTTGTTGGA ATGTCCTCAA CCCTAGTATC
 651 TTCTGTTCCC ACAGCAGCTG TGCCCCCCT GCCTAAACGGG GCTCCCCCTG
 701 TTATACAAAC TCTGCCCTGCA TTGCTCATC CTGGAGCCAC ATGGCAAAG
 751 AGTTCTCTT TTAGTAGATC TGGTCCAGGG TCACAACCAA ACACTAAATT
 801 ACAAAAGGCA CAGTCATTG ATGTCGCAG TGTCGCCACCA GTGGCAGAGT
 851 GGGCTGTTOC TCAGTCATCA AGACTGAAAT ACAGGCAATT ATTCAATAGT
 901 CATGACAAAA CTATGAGTGG ACACCTAACCA GGTCCCCAAG CAAGAACTAT
 951 TCTTATGCAG TCAAGTTAC CACAGGCCTCA GCTGGCTTCA ATATGGAATC
 1001 TTCTGACAT TGATCAAGAT GGAAACCTTA CACCGAGGAA ATTATCTG

Figure 4

1051 GCAATGCACC TCAATTGATGT AGCTATGCTC GGCCAACAC TGCCACCTGT
1101 CCTGGCTCCA GAATAACATTC CACCTTCTTT TAGAAGAGTT CGATCTGGCA
1151 GTGGTATATC TGTCATAAGC TCAACATCTG TAGATCAGAG GCTACCAGAG
1201 GAACCAGTTT TAGAAGATGA ACAACAACAA TTAGAAAAGA AATTACCTGT
1251 AACGTTTGAA GATAAGAACG GGGAGAACTT TGAACGTGGC AACCTGGAAC
1301 TGGAGAAACG AAGGCAAGCT CTCCCTGGAAC ACCAGCGCAA GGAGCAGGAG
1351 CGCCCTGGCCC AGCTGGAGCG GGOGGAGCAG GAGAGGAAGG AGCGTGAGCG
1401 CCAGGAGCAA GAGCGCAAAA GACAACCTGGA ACTGGAGAAG CAACTGGAAA
1451 AGCAGCGGGA GCTAGAACCG CAGAGAGAGG AGGAGAGGAG GAAAGAAATT
1501 GAGAGGAGAG AGGCTGCAAA ACGGGAACCTT GAAAGGCAAC GACAACCTGA
1551 GTGGGAACCG AATCGAAGGC AAGAACTACT AAATCAAAGA AACAAAGAAC
1601 AAGAGGACAT AGITGTACTG AAAGCAAAGA AAAAGACTTT CGAATTGAA
1651 TTAGAAGCTC TAAAATGATAA AAAGCATCAA CTAGAAGGGA AACCTCAAGA
1701 TATCAGATGT CGATTGACCA CCCAAAGGCA AGAAATIGAG AGCACAAACA
1751 AAATCTAGAGA GTTGTAGAATT GCGAAATCA CCCATCTACA GCAACAAITA
1801 CAGGAATCTC AGCAAATGCT TGGAAAGACTT ATTCAGAAA AACAGATACT
1851 CAATGACCA TTTAAACAAG TTICAGCAGAA CAGTTTGCAC AGAGATTAC
1901 TTTGTTACACT TAAAAGAGCC TTAGAAGCAA AAGAACTAGC TOGGCAGCAC
1951 CTACGAGACC AACTGGATGA AGTGGAGAAA GAAACTAGAT CAAAACCTACA
2001 GGAGATTGAT ATTTCAATA ATCAGCTGAA GGAACTAAGA GAAATACACA
2051 ATAAGCAACA ACTCCAGAAG CAAAAGTCCA TGGAGGCTGA ACGACTGAAA
2101 CAGAAAGAAC AAGAACCCAAA GATCATAGAA TTAGAAAAAC AAAAAGAAGA
2151 AGCCCCAAAGA CGAGCTCAGG AAAGGGACAA GCAGTGGCTG GAGCATGTC
2201 AGCAGGAGGA CGAGGCATCAG AGACCAAGAA AACTCCACGA AGAGGAAAAA
2251 CTGAAAAGGG AGGAGAGTGT CAAAAAGAAC GATGGCGAGG AAAAAGGCAA

Figure 4

2301 ACAGGAAGCA CAAGACAAGC TGGGTGGCT TTICCATCAA CACCAAGAAC
2351 CAGCTAAGCC AGCTGTCAG GCACCCCTGGT CCACTGCAGA AAAAGGTCCA
2401 CTTACCAATT CTGCACAGGA AAAATGAAAAA GTGGTGTTATT ACGGGCACT
2451 GTACCCCTTT GAATCCAGAA GGCATGATGA AATCACTATC CAGCCAGGAG
2501 ACATAGTCAT GGTTGGATGAA AGCCAAACTG GAGAACCGG CTGGCTTGGA
2551 CGAGAATTAA AAGGAAAGAC AGGGTGGITC CCTGCAAAC ACTGAGAGAA
2601 AATCCCAGAA AATGAGGTTG CGCTCCAGT GAAACCAGTG ACTGATTCAA
2651 CATCTGCCOC TGCCCCAAA CTGGCCTTGC GTGAGACCCC CGCCCCTTTG
2701 GCAGTAACCT CTTCAGAGCC CTCCACGACC CCTAATAACT GGGCGACTT
2751 CAGCTCCACG TGGCCACCA GCACGAATGA GAAACCAGAA ACGGATAACT
2801 GGGATGCATG GGCAGCCAG CCTCTCTCA CGTTCCAAG TGCGGGCCAG
2851 TTAAGGCAGA GGTCGCCCTT TACTCCAGCC ACGGCCACTG GCTCTCCOC
2901 GTCTCTGTG CTAGGCCAGG GTGAAAAGGT GGACGGGCTA CAAGCTCAAG
2951 CCTATATATCC TTGGAGAGCC AAAAAAGACA ACCACTTAAA TTTTAACAAA
3001 AATGATGTC TCAOCGTCT GGAACAGCAA GACATGIGGT GGTTCGGAGA
3051 AGTCAAGGT CAGAAGGGIT GGTTCGCCAA GTCTTACGGG AAACTCATT
3101 CAGGGCCCAT AAGGAAGTCT ACAAGCATGG ATTCTGGITC TTCAAGAGT
3151 CCTGCTAGTC TAAAGOGAGT AGCCTCTCCA GCAGCCAAGC CGGTGTTTC
3201 GGGAGAAGAA ATTGCCAGG TTATTGCTC ATACACGCC ACCGGCOOCG
3251 AGCAGCTCAC TCTGGCCCT GGTCAAGCTGA TTTTGATCOG AAAAAAGAAC
3301 CCAGGTGGAT GGTGGGAAGG AGAGCTGCAA GCAAGTGGGA AAAAGGCCA
3351 GATAGGCTGG TTCCAGCTA ATTATGAAAA GCTTCTAAGC CCTGGGACGA
3401 GCAAAATCAC TCCAACAGAG CCACCTAAGT CAACAGCAATT AGCGGCAGTG
3451 TGCCAGGIGA TTGGGATGTA CGACTACACC GCGCAGAATG ACGATGAGCT

Figure 4

3501 GGCCITCAAC AAGGCCAGA TCATCAACGT CCTCAACAAG GAGGACOCTG
3551 ACTGGTGGAA AGGAGAACGTC AATGGACAAG TGGGGCTCTT CCCATCCAAT
3601 TATGTGAAGC TGACCACAGA CATGGACCCA AGCCAGCAAT GAATCATAATG
3651 TTGTCCATCC CCCCCCICAGG CTTGAAAGTC CTCAAAGAGA CCCACTATCC
3701 CATATCACTG CCCAGAGGGA TGATGGAGA TGCAGCCITG ATCATGTGAC
3751 TTCCAGCATG ATCACCTACT GCCTTCCTGAG TAGAAGAACT CACTGCAGAG
3801 CAGTTTACCT CATTTTACCT TAGTTGCATG TGATGCCAT TTTTGTGTTA
3851 TTACTTGCAG AGATAGGAGC AAAAATTACA AAAACACACA GGGTAGTGGG
3901 TCCTTTGTG GCCTTCCTAG TTACTCAAT TGACTTCCC CCACCTTTGC
3951 ACAGGTGCTT TCAATAGTTT TAAAATTATT TTTAAATATA TATTTTAGCT
4001 TTTTAATAAA CAAAATAAT AAATGACTTC TTTGCTATT TGGTTTGCA
4051 AAAAGACCCA CTATCAAGGA ATGCTGGATG TGCTATTAAA AATIGITCCA
4101 AATGTCATAA AATCTGAGAC TIGATGTTATT TTTTCACTTT GTCCAGTGT
4151 ACCAACTAAA TTGCTGCAGT TTGGGGCTT TCCOCCTTAC CATAGAAGTG
4201 CAGAGGAGTT CAGTATCTCT GTTTAAAGA CGTATAGAAT GAGCCAAATT
4251 AAAGCGAAGG TGATTGIGCT TGTTTGTGIG TATCAGCTGT ACCTTGTGTA
4301 GCAATGTAATA CATCCCTGTC ATAAGAAATT AGTTCTTCC ATGGCAAAGC
4351 TAATACCTTG TACGATGTC TAATCATATT GCATTTAAATT TTAATTGCA
4401 aCAGTGACCT TGTAGCCACA TGAGAAAGCA CTCTGIGTTT TTGTTGGTC
4451 TCAGATTAT CTGGTTGAGT TGGTGTGTTG TTGGGGTTT TTAATTTCG
4501 GIGTTTGCAT AGCATAAAAT CAGTAGACAA CACCACTGAG GTGGTACGA
4551 TCAACGATAT CCACAGTCIC TTTTACGTCCT CTGTTACATG AAGTTTTATT
4601 CCAGTTACTT TTCACTGAAAT GACCTATT TAAACAAGTAA TTTTCCTGAC
4651 AAGAAAGAT GTATAGAAGT CTCCCTGCAA TTAAATTCCA ATGTTTACAT
4701 TTTTTAACTA GGACTGTGGA ATTCTACAG ATTAATATGA AATGGAGCTC

4751 ATGGTCCGTT TGIGTGTTAG ATATGCTGTA GCTGAAGCCC TGTTTGTCTT
4801 TAAACACTA GTTGGAAAGCT CTCATAAAA ATGCCCTGCTG CTCACAGCAC
4851 AGAAAAATGGG GCAGGGGGAG CCTCAAGCAC AATCTAGCTG TCCCTCTAAA
4901 GACTCTGTAA TGCTCAATCC CCTTGGGTTC TCCCCGGGCT GTGGGGAGGC
4951 TGTGCTGGTG GTGGAGTAGA GGTCCTTTTC CTTCAAATG GTGGAGAGAG
5001 AGAGGACCTT TCCCTCTTGT TCAGTTGCAA TTCACTATTTC TCACGGATAT
5051 GAATGTAAAA TATATAAATA TATAAACCTG AGGATTTAAC AAATGTAAAA
5101 CAACCTTTTG AATTAGTTCC GAGTATAGAT AATTAAATTTC TTAAAACAAA
5151 AGTAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAGCTGAC GGGCGCG

DRAFT - NOT FOR CITATION

Figure 4

SH3D1A Translated Protein Sequence:

1 MAQFPTPPGG SLDIWAITVE ERAKHDQQFH SLKPISGFIT GDQARNFFFO
51 SGLPQPVLAQ IWALADMNND GRMDQVEFSI AMKLIKLKQ GYOLPSALPP
101 VMKQQPVAIS SAPAFGMGGI ASMPLTAVA PVPMSIPVV QMSPTILVSSV
151 PTAAVPPLAN GAPPVIQPLP AFAHPAATLP KSSSFSRSGP GSQNLNIKLOK
201 AQSFIDVASVP PVAEWAVPQS SRLKYRQLFN SHDKIMSGHL TGRQARTILM
251 QSSLPOAQLA SIWNLSDIDQ DGKLTAEEFI LAMHLIDVAM SGQPLPPVLP
301 PEYIPPSFRR VRSGSGGISVI SSTSVDQRLP EEPVLEDEQQ QLEKKLPVIF
351 EDKKRENPER GNLELEKRRQ ALLEQORKEQ ERLAQLERAE QERKERERQE
401 QERKROLELE KOLEKORELE ROREEERRKE IERREAAKRE LEROROLEWE
451 RNRROELLNQ RNKEQEDIVV LKAKKKTLEF EFLALNDKKH OLEGKLODIR
501 CRLTTQRQEI ESTINKSRELR IAETTHLQQQ LQESQOMLGR LIPEKQILND
551 QLKQVOONSL HRDSLWILKR ALEAKELARQ HLRDQOLDEVE KETRSKLOEI
601 DIFNNOLKEL REIHNKQOLO KQKSMEAERL KQKEQERKII ELEKOKEEAQ

09/720934

651 RRAQERDKOW LEHVQQEDEH QRPRKLHEEE KLKREESVKK KDGEEKGQE
701 AODKLGRLFH QHQEPAKPAV QAPWSTAEGP PLTISAQENV KVYYYRALYP
751 FESRSHDEIT IQPGDIVMVD ESOIGEPGWL GGELKGKIGW FPANYAEKIP
801 ENEVPAPVKP VIDSTSAPAP KLALRETPAP LAVTSSEPST TPNNWADFSS
851 TWPTSTINEKP EIDNWDAWAA QPSLTIVPSAG QLRQRSAFTP ATATGSSPSP
901 VLGCQGEKVEG LQAQALYPWR AKKDNHLPN KNDVITVLEQ QDMWWFGEVQ
951 GOKGWFPKSY VKLISGPIRK STSMDSGSSE SPASLKRVAS PAAKPVVSGE
1001 EIAQVIASYT ATGPEDQTLA PGQLILIRKK NPGGWWEIGL QARGKKRQIG
1051 WFPANYVKLL SPGTTSKITPT EPPKSTALAA VCGVIGMYDY TAQNDDELAF
1101 NKGQIINVLN KEDPDWWKGE VNGQVGLFPS NYVKLTIDMD PSQ

09/720934

WO 99/53062

11 / 30

PCT/US99/08371

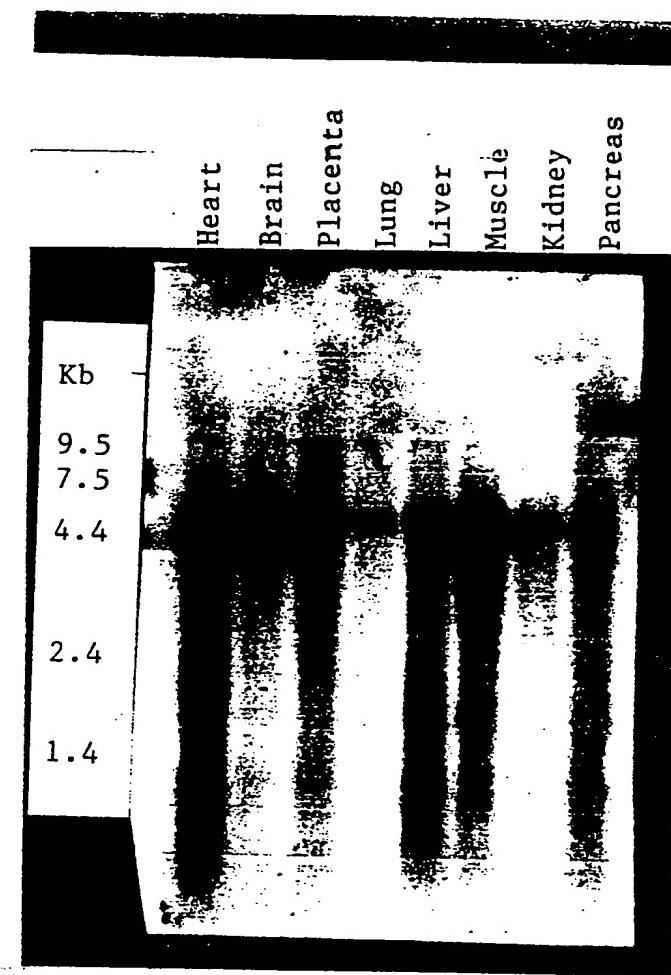


Figure 6

23220-1-001 PCT (Sheet 12 of 30)

Summary of cDNAs isolated

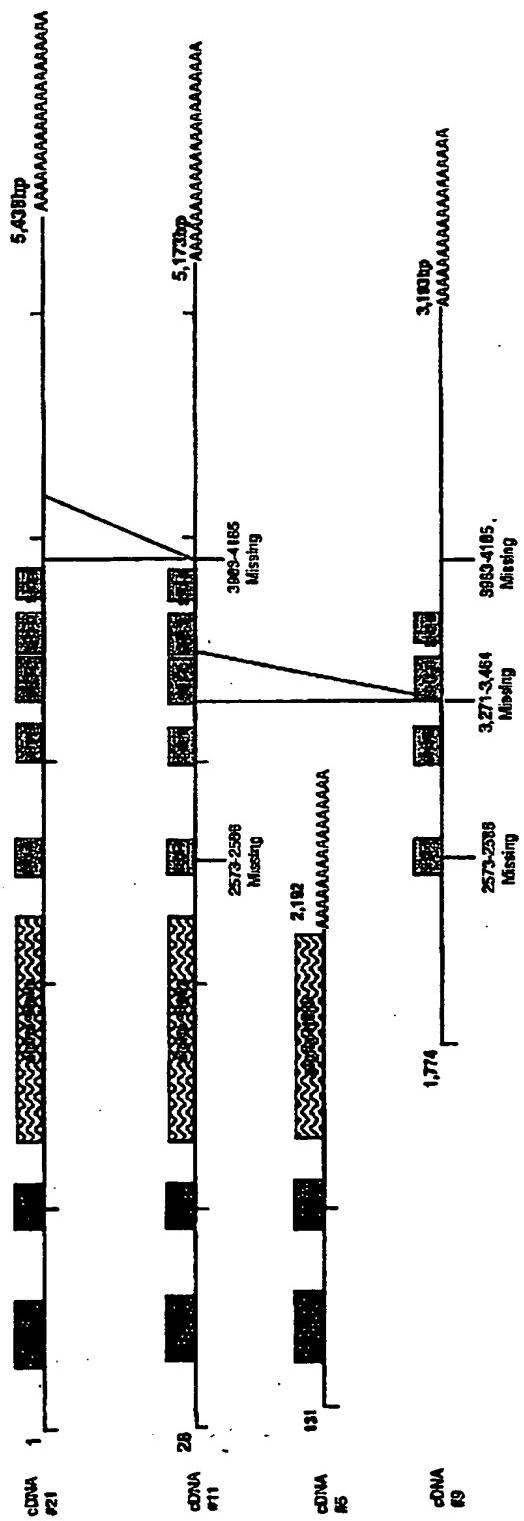


Figure 7

1 GCACGGAGG GAGCGAAGGA GGTAGAGAAG AGTGGAGGCG CCAGGGGAGG
51 GAGCGTAGCT TGGTTGCTCC GTAGTACGGC GGCTCGCGAG GAAGAATCCC
101 GAGCGGGCTC CGGGACGGAC AGAGAGGCAGG GCAGGGATGG TGTGCGGGGC
151 TGCAGCTCCT CGTCCTCC CAGCGGCCG TGAGCGGCAC TGATTGTCC
201 CTGGGGCGGC AGCGCGGACC CGCCCGGAGA TGAGGCCTCG ATTAGCAAGG
251 TAAAAGTAAC AGAACCATGG CTCAGTTCC AACACCTTT GGTGGCAGCC
301 TGGATATCTG GGCCATAACT GTAGAGGAAA GAGCGAAGCA TGATCAGCAG
351 TTCCATAGTT TAAAGCCAAT ATCTGGATT ATTACTGGTG ATCAAGCTAG
401 AAACTTTTT TTTCAATCTG GGTTACCTCA ACCTGTTTA GCACAGATAT
451 GGGCACTAGC TGACATGAAT AATGATGGAA GAATGGATCA AGTGGAGTTT
501 TCCATAGCTA TGAAACTTAT CAAACTGAAG CTACAAGGAT ATCAGCTACC
551 CTCTGCACCT CCCCTGTCA TGAAACAGCA ACCAGTTGCT ATTCTAGCG
601 CACCAGCATT TGGTATGGGA GGTATGCCCA GCATGCCACC GCTTACAGCT
651 GTTGCCTCCAG TGCCATGGG ATCCATTCCA GTTGTGGAA TGTCTCCAAC
701 CCTAGTATCT TCTGTTCCCA CAGCAGCTGT GCCCCCCCTG GCTAACGGGG
751 CTCCCCCTGT TATACAACCT CTGCCTGCAT TTGCTCATCC TGCAGCCACA
801 TTGCCAAAGA GTTCTCCCT TAGTAGATCT GGTCAGGGT CACAACAAA
851 CACTAAATTA CAAAAGGCAC AGTCATTGA TGTCAGGAGT GTCCCACCAAG
901 TGGCAGAGTG GGCTGTTCCCT CAGTCATCAA GACTGAAATA CAGGCAATTA
951 TTCAATAGTC ATGACAAAAC TATGAGTGGGA CACTTAACAG GTCCCCAAGC
1001 AAGAACTATT CTTATGCAGT CAAGTTACC ACAGGCTCAG CTGGCTTCAA
1051 TATGGAATCT TTCTGACATT GATCAAGATG GAAAACCTAC AGCAGAGGAA
1101 TTTATCCTGG CAATGCACCT CATTGATGTA GCTATGTCTG GCCAACCACT
1151 GCCACCTGTC CTGCCTCCAG AATACATTCC ACCTTCTTT AGAAGAGTTC
1201 GATCTGGCAG TGGTATATCT GTCATAAGCT CAACATCTGT AGATCAGAGG
1251 CTACCAGAGG AACCAGTTT AGAAGATGAA CAACAACAAT TAGAAAAGAA
1301 ATTACCTGTA ACCTTGAAG ATAAGAAGCG GGAGAACTTT GAACGTGGCA
1351 ACCTGGAACT GGAGAAACGA AGGCAAGCTC TCCTGGAACA GCAGCGCAAG
1401 GAGCAGGAGC GCCTGGCCCA GCTGGAGCGG GCGGAGCAGG AGAGGAAGGA
1451 GCGTGAGCGC CAGGAGCAAG AGCGAAAAG ACAACTGGAA CTGGAGAAGC
1501 AACTGGAAAA GCAGCGGGAG CTAGAACGGC AGAGAGAGGA GGAGAGGAGG
1551 AAAGAAATTG AGAGGCAGA GGCTGAAAAA CCGGAACCTTG AAAGGCAACG
1601 ACAACTTGAG TGGGAACCGA ATCGAAGGCA AGAACTACTA AATCAAAGAA
1651 ACAAAAGAACAGAGGACATA GTTGTACTGA AAGCAAAGAA AAAGACTTTG
1701 GAATTGAAAT TAGAAGCTCT AAATGATAAA AAGCATCAC TAGAAGGGAA
1751 ACTTCAAGAT ATCAGATGTC GATTGACCAC CCAAAGGCAA GAAATTGAGA
1801 GCACAAACAA ATCTAGAGAG TTGAGAATTG CCGAAATCAC CCATCTACAG
1851 CAACAATTAC AGGAATCTCA GCAAATGCTT GGAAGACTTA TTCCAGAAAA
1901 ACAGATACTC AATGACCAAT TAAAACAAGT TCAGCAGAAC AGTTGCACA
1951 GAGATTCACT TGTTACACTT AAAAGAGCCT TAGAAGCAAA AGAACTAGCT
2001 CGGCAGCACC TACGAGACCA ACTGGATGAA GTGGAGAAAG AACTAGATC
2051 AAAACTACAG GAGATTGATA TTTCAATAA TCAGCTGAAG GAACTAAGAG
2101 AAATACACAA TAAGCAACAA CTCCAGAACG AAAAGTCCAT GGAGGCTGAA

Figure 8

2151 CGACTGAAAC AGAAAGAAC AGAACGAAAG ATCATAGAAT TAGAAAAACA
2201 AAAAGAAGAA GCCCAAAGAC GAGCTCAGGA AAGGGACAAG CAGTGGCTGG
2251 AGCATGTGCA GCAGGAGGAC GAGCATCAGA GACCAAGAAA ACTCCACGAA
2301 GAGGAAAAAC TGAAAAGGGA GGAGAGTGTG AAAAAGAAGG ATGGCGAGGA
2351 AAAAGGCAAA CAGGAAGCAC AAGACAAGCT GGGTCGGCTT TTCCATCAAC
2401 ACCAAGAACCC AGCTAAGCCA GCTGTCCAGG CACCCTGGTC CACTGCAGAA
2451 AAAGGTCCAC TTACCATTTG TGACAGGAA AATGTAAAAG TGGTGTATTA
2501 CGGGGCAC TG TACCCCTTG AATCCAGAAG CCATGATGAA ATCACTATCC
2551 AGCCAGGAGA CATAGTCATG GTTAAAGGGG ATGGGTGGA TGAAAGCCAA
2601 ACTGGAGAAC CCGGCTGGCT TGGAGGAGAA TTAAAAGGAA AGACAGGGTG
2651 GTTCCCTGCA AACTATGCAG AGAAAATCCC AGAAAATGAG GTTCCCGCTC
2701 CAGTGAAACC AGTGAATGAT TCAACATCTG CCCCTGCC CAAACTGGCC
2751 TTGCGTGAGA CCCCCCCCCC TTTGGCAGTA ACCTCTTCAG AGCCCTCCAC
2801 GACCCCTAAT AACTGGGCCG ACTTCAGCTC CACGTGGCCC ACCAGCACGA
2851 ATGAGAAACC AGAAACGGAT AACTGGGATG CATGGGCAGC CCAGCCCTCT
2901 CTCACCGTTC CAAGTGCAGG CCAGTTAAGG CAGAGGTCCG CCTTTACTCC
2951 AGCCACGGCC ACTGGCTCCT CCCCCGCTCC TGTGCTAGGC CAGGGTGAAA
3001 AGGTGGAGGG GCTACAAGCT CAAGCCCTAT ATCCCTTGGAG AGCCAAAAAA
3051 GACAACCACT TAAATTAA CAAAAATGAT GTCATCACCG TCCTGGAAACA
3101 GCAAGACATG TGGTGGTTG GAGAAGTTCA AGTCAGAAG GTTGGTTCC
3151 CCAAGTCTTA CGTAAACTC ATTTCAGGGC CCATAAGGAA GTCTACAAGC
3201 ATGGATTCTG GTTCTTCAGA GAGTCCTGCT AGTCTAAAGC GAGTAGCCTC
3251 TCCAGCAGCC AAGCCGGTCA TTTCGGGAGA AGAATTATT GCCATGTACA
3301 CTTACGAGAG TTCTGAGCAA GGAGATTAA CCTTCAGCA AGGGGATGTG
3351 ATTTGGTTA CCAAGAAAAGA TGGTGAATGG TGGACAGGAA CAGTGGCGA
3401 CAAGGCCCGA GTCTCCCTT CTAACTATGT GAGGCTTAAA GATTAGAGG
3451 GCTCTGGAAC TGCTGGAAA ACAGGGAGTT TAGGAAAAAA ACCTGAAATT
3501 GCCCAGGTTA TTGCTCATA CACCGCCACC GGCCCCGAGC AGCTCACTCT
3551 CGCCCTGGT CAGCTGATTT TGATCCGAAA AAAGAACCC GGTGGATGGT
3601 GGGAGGAGA GCTGCAAGCA CGTGGGAAAA AGCGCCAGAT AGGCTGGTTC
3651 CCAGCTAATT ATGTAAGCT TCAAGCCCT GGGACGAGCA AAATCACTCC
3701 AACAGAGCCA CCTAAGTCAA CAGCATTAGC GGCAGTGTG CAGGTGATTG
3751 GGATGTACGA CTACACCGCG CAGAATGACG ATGAGCTGGC CTTCAACAAG
3801 GGCCAGATCA TCAACGTCTT CAACAAGGAG GACCCCTGACT GGTGGAAAGG
3851 AGAAAGTCAAT GGACAAGTGG GGCTCTTCCC ATCCAATTAT GTGAAGCTGA
3901 CCACAGACAT GGACCAAGC CAGCAATGAA TCATATGTTG TCCATCCCCC
3951 CCTCAGGCTT GAAAGTCCTC AAAGAGACCC ACTATCCCAT ATCACTGCC
4001 AGAGGGATGA TGGGAGATGC AGCCTGATC ATGTGAATTC CAGCATGATC
4051 ACCTACTGCC TTCTGAGTAG AAGAACTCAC TGCAAGAGCAAG TTTACCTCAT
4101 TTTACCTTAG TTGCAATGTGA TCGCAATGTT TGAGTTATTA CTTGCAGAGA
4151 TAGGAGCAAA ATTACAAAA ACACACAGGG TAGTGGGTCC TTTTGTGGCT
4201 TTCTCTAGTTA CTCAAATTGA CTTTCCCCCA CCTTTGCACA GGTGCTTCA
4251 ATAGTTTAA ATTATTTT AAATATATAT TTAGCTTTT TAATAAACAA
4301 AATAAATAAA TGACTCTTT GCTATTTGG TTTGCAAAA AGACCCACTA
4351 TCAAGGAATG CTGCATGTGC TATTAATAT TGTTCCTAAAT GTCCATAAAT

Figure 8

09/720934

WO 99/53062

PCT/US99/08371

15 / 30

4401 CTGAGACTTG ATGTATTTT TCA TTTTGTCA CAGTGTACCA AACTAAATTG
4451 TGCAGTTGG GGCTTTCCC CCTTACCATA GAAGTGCAGA GGAGTTCA GT
4501 ATCTCTGTT TAAAGACGTA TAGAATGAGC CCAATTAAAG CGAAGGTGTT
4551 TGTGTTGTT TGTGTGTATC AGCTGTACCT TGTGAGCAT GTAATACATC
4601 CTGTACATAA GAAATTAGTT CTTCCATGG CAAAGCTATT ACCTTGTACG
4651 ATGCTCTAAT CATATTGCAT TAAATTTAT TTTGCACAGT GACCTTGTAG
4701 CCACATGAGA AAGCACTCTG TGTTTTGTT CGGTCTCAGA TTTATCTGGT
4751 TGAGTTGGTG TTTGTTGG GTTTTTAAT TTGCGTGTGTT TGCATAGCAT
4801 AAAATCAGTA GACAACACCA CTGAGGTGCGT TACGATCAAC GATATCCACA
4851 GTCTCTTTT AGTCTCTGTT ACATGAAGTT TTATTCCAGT TACTTTTCAT
4901 GGAATGACCT ATTTGAACA AGTAATTTC TTGACAAGAA AGAATGTATA
4951 GAAGTCTCCC TGCAATTAAAT TTCCAATGTT TACATTTTT AACTAGACTG
5001 TGGAATTCT ACAGATTAAT ATGAAATGGA GCTCATGGTC CGTTTGTG
5051 TTAGATATGC TGTAGCTGAA GCCCTGTTG TCTTTAAC ACTAGTTGGA
5101 AGCTCTCAAT AAAATGCCT GCTGCTCACA GCACAGAAAA TGGGGCAGGG
5151 GGAGCCTCAA GCACAATCTA GCTGCTCTCC TAAAGACTCT GTAATGCTCA
5201 CTCCCCCTCGC GTTCTCCCGG CGCTGTCGGG AGGCTGTGCT GGTGGTCGTG
5251 TAGAGGTCCCT TCTCCTTCA CATGGTGAG AGAGCGAGGA CCTCTCCTCC
5301 TCGTTCAGTT GCACTTCAGT ATTTCACGG ATATGAATGT AAAATATATA
5351 AATATATAAAA CCTGCGGCTT TAACAACGT AATACAACCT TTTGAATTAG
5401 TTCCGTGTAT AGATAATTAA ATTCTTCATA CAAAAGTTAA AAAAAAAAAA
5451 AAAAAAAA

Figure 8

#21 translated protein sequence:

1 MAQFPTPF GG SLDIWAITVE ERAKHDQQFH SLKPISGFIT GDQARNFFFQ
51 SGLPQPVLAQ IWALADMNND GRMDQVEFSI AMKLILKLQ GYQLPSALPP
101 VMKQQPVAIS SAPAFGMGGI ASMPPLTAVA PVPMSIPVV GMSPTLVSSV
151 PTAAVPPLAN GAPPVIQPLP AFAHPAATLP KSSSFSRSGP GSQNLNTKLQK
201 AQSFDVASVP PVAEWAVPQS SRLKYRQLFN SHDKTMSGHL TGPQARTILM
251 QSSLPQAQLA SIWNLSDIDQ DGKLTAEEFI LAMHLDIVAM SGQPLPPVLP
301 PEYIPPSFRR VRSGSGVISI SSTSVDQR LP EEPVLEDEQQ QLEKKLPVTF
351 EDKKRENFER GNLELEKRRQ ALLEQQRKEQ ERLAQLERAQ QERKERERQE
401 QERKRQLELE KQLEKQRELE RQREEERRKE IERREAQKRE LERQRQLEWE
451 RNRRQELLNQ RNKEQEDIVV LKAKKKTLEF ELEALNDKKH QLEGKLQDIR
501 CRLTTQRQEI ESTNKSRELR IAEITHLQQQ LQESQQMLGR LIPEKQILND
551 QLKQVQQNSL HRDSLVTLKR ALEAKELARQ HLRDQLDEVE KETRSKLQEI
601 DIFNNQLKEL REIHNKQQLQ KQKSMEAERL KQKEQERKII ELEKQKEEAQ
651 RRAQERDKQW LEHVQQEDEH QRPRKLHEEE KLKREESVKK KDGEEKGKQE
701 AQDKLGRLFH QHQEPAKPAV QAPWSTAEG PLTISAQENV KVYYYRALYP
751 FESRSRSHDEIT IQPGDIVMVK GEWVDESQTG EPGWLGGELK GKTGWFANY
801 AEKIPENEVP APVKPVTDST SAPAPKLALR ETPAPLAVTS SEPSTTPNNW
851 ADFSSTWPTS TNEKPETDNW DAWAAQPSLT VPSAGQLRQR SAFTPATATG
901 SSPSPVLGQG EKVEGLQAQA LYPWRACKDN HLNFNKNDVI TVLEQQDMWW
951 FGEVQQQKGW FPKSYVKLIS GPIRKSTSMD SGSSESPASL KRVASPAAKP
1001 VVSGEEFIAM YTYESSEQGD LTFQQGDVIL VTKKDGDWWT GTVGDKAGVF
1051 PSNYVRLKDS EGSGTAGKTG SLGKKPEIAQ VIASYTATGP EQLTLAPGQL
1101 ILIRKKNP GG WWEHELQARG KKRQIGWPA NYVKLLSPGT SKITPTTEPPK
1151 STALAAVCQV IGYMYDYTAQN DDELAFNKGQ IINVLNKEDP DWWKGEVNGQ
1201 VGLFPSNYVK LTTDMDPSSQ *

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30

Figure 9

Whole protein sequence

1 TRGSEGGREE WRRQGRERSL VAP*YGGSRG RIPSGLRDGQ RGGRGWCAGL
 51 RLLRPSQRRV SGTDLSLGRQ RGPARR*GVD *QGKSNRTMA QFPTPFGGSL
 101 DIWAITVEER AKHDQQFHSL KPIISGFITGD QARNFFFQSG LPQPVLAQIW
 151 ALADMNNNDGR MDQVEFSIAM KLIKLLQGY QLPSALPPVM KQQPVAISSA
 201 PAFGMGGIAT MPPLTAVAPV PMGSIPVVGGM SPTLVSSVPT AAVPPLANGA
 251 PPVIQPLPAF AHPAATLPKS SSFSRSRGPGS QLNTKLQKAQ SFDVASVPPV
 301 AEWAVPQSSR LKYRQLFNH DKTMSGHLTG PQARTILMQS SLPQAQLASI
 351 WLNSDIDQDG KLTAEEFILA MHLIDVAMSG QPLPPVLPE YIPPSFRRVR
 401 SGSGISVISS TSVVDQRLPEE PVLEDEQQQL EKKLPVTFED KKRENFERGN
 451 LELEKRRQAL LEQQRKEQER LAQLERAEQE RKERERQEQE RKRQLELEKQ
 501 LEKQRELERQ REEERRKEIE RREAAKRELE RQRQLEWERN RRQELLNQRN
 551 KEQEDIVVLK AKKKTLEFEL EALNDKKHQL EGKLQDIRCR LTTQRQEIES
 601 TNKSRELRIA EITHLQQQLQ ESQQMLGRLL PEKQILNDQL KQVQQNSLHR
 651 DSLVTLKRAL EAKELARQHL RDQLDEVEKE TRSKLQEIDI FNNQLKELRE
 701 IHNKQQLQKQ KSMEAERLKQ KEQERKIEL EKQKEEAQRR AQERDKQWLE
 751 HVQQEDEHQR PRKLHEEEKL KREESVKKKD GEEKGKQEAQ DKLGRLFHQH
 801 QEPAKPAVQA PWSTAEGGPL TISAQENVKV VYYRALYPFE SRSHDEITIQ
 851 PGDIVMVKG EWVDESQTGE GWLGGEKGK TGWFPANYAE KIPENEVPAP
 901 VKPVTDSTSA PAPKLALRET PAPLAUTSSE PSTTPNNWAD FSSTWPTSTN
 951 EKPETDNWDA WAAQPSLTVP SAGQLRQRSA FTPATATGSS PSPVLGQGEK
 1001 VEGLQAQALY PWRAKKDNHL NFNKNDVITV LEQQDMWWFG EVQQQKGWFP
 1051 KSYVKLISGP IRKSTSMDSG SSESASPASLKR VASPAAKPVV SGEEFIAMYT
 1101 YESSEQGDLT FQQGDVILVT KKDGDWWTGT VGDKAGVFFPS NYVRLKDSEG
 1151 SGTAGKTGSL GKKPEIAQVI ASYTATGPEQ LT LAPGQLIL IRKKNPAGWW
 1201 EGELQARGKK RQIGWFPPANY VKLLSPGTSK ITPTEPPKST ALAAVCQVIG
 1251 MYDYTAQNDD ELAFNKGQII NVLNKEDPDW WKGEVNGQVG LFPSNYVKLT
 1301 TDMDPSQQ*I ICCPSPPQA* KSSKRPTIY HCPEG*WEMQ P*SCDFQHDH
 1351 LLPSE*KNSL QSSLPHFTLV ACDRNV*VIT CRDRSKNYKN TQGSGSFCCF
 1401 PSYSN*LSPT FAQVLSIVLK LFLNIYFSFL INKINK*LLC YFGFAKRPTI
 1451 KECCMCY*KL FQMSINLRD VFHFVQCYQ LNCAVWGFSP LP*KCRGVQY
 1501 LCFKDV*NEP N*SEGVACL CVSAVPC*AC NT SCT*EISS FHGKAITLYD
 1551 ALIILHLILF CTVTL*PHEK ALCVFVRSQI YLVELVFCLG FLILRVCIA*
 1601 NQ*TTPRLSL RSTISTVSF* SLLHEVLFQL LFME*PILNK *FS*QERMYR
 1651 SLPAINFQCL HFLTRLWNFY RLI*NGAHGP FVC*ICCS*S PVCLLNTSWK
 1701 LSIKMPAAHS TENGAGGASS TI*LSS*RLC NAHSPRVLPA LSGGCAGGRV
 1751 EVLLLSHGAE SEDLSSSSFSC TSVFSRI*M* NI*IYKPAAL TTVIQPFEV
 1801 PCIDN*ILHT KVKKKKKK

Figure 9

1 AGAGTGGAGG CGCCAGGGGA GGGAGCGTAG CTTGGTTGCT CCGTAGTACG
51 GCGGCTCGCG AGGAAGAAC CCGAGCGGGC TCCGGGACGG ACAGAGAGGC
101 GGGCGGGAT GGTGTGCGGG GCTGCGGCTC CTGCGTCCCT CCCAGCGCG
151 CGTGAGCGGC ACTGATTGT CCTGGGGCG GCAGCGCGGA CCCGCCCGGA
201 GATGAGGCCT CGATTAGCAA GTAAAAGTA ACAGAACCAT GGCTCAGTT
251 CCAACACCTT TTGGTGGCAG CCTGGATATC TGGGCCATAA CTGTAGAGGA
301 AAGAGCGAAG CATGATCAGC AGTTCCATAG TTTAAAGCCA ATATCTGGAT
351 TCATTACTGG TGATCAAGCT AGAAACTTIT TTTTCAATC TGGGTTACCT
401 CAACCTGTT TAGCACAGAT ATGGGCACTA GCTGACATGA ATAATGATGG
451 AAGAATGGAT CAAGTGGAGT TTTCCATAGC TATGAAACTT ATCAAACCTGA
501 AGCTACAAGG ATATCAGCTA CCCTCTGCAC TTCCCCCTGT CATGAAACAG
551 CAACCAAGTT CTATTTCTAG CGCACCCAGCA TTTGGTATGG GAGGTATCGC
601 CAGCATGCCA CCGCTTACAG CTGTTGCTCC AGTCCAATG GGATCCATT
651 CAGTTGTTGG AATGTCTCCA ACCCTAGTAT CTTCTGTTCC CACAGCAGCT
701 GTGCCCCCCC TGGCTAACGG GGCTCCCCCT GTTATACAAC CTCTGCCTGC
751 ATTTGCTCAT CCTGCAGCCA CATTGCCAAA GAGTTCTTCC TTTAGTAGAT
801 CTGGTCCAGG GTCACAACTA AACACTAAAT TACAAAAGGC ACAGTCATT
851 GATGTGGCCA GTGTCCCACC AGTGGCAGAG TGGGCTGTT CTCAGTCATC
901 AAGACTGAAA TACAGGCAAT TATTCAATAG TCATGACAAA ACTATGAGTG
951 GACACTTAAC AGGTCCCCAA GCAAGAAC TA TTCTTATGCA GTCAAGTTA
1001 CCACAGGCTC AGCTGGCTTC AATATGGAAT CTTCTGACA TTGATCAAGA
1051 TGGAAAACCTT ACAGCAGAGG AATTATCCT GGCAATGCAC CTCATTGATG
1101 TAGCTATGTC TGGCCAACCA CTGCCACCTG TCCTGCCTCC AGAATACATT
1151 CCACCTTCTT TTAGAAGAGT TCGATCTGGC AGTGGTATAT CTGTCTATAAG
1201 CTCAACATCT GTAGATCAGA GGCTACCAGA GGAACCAGTT TTAGAAGATG
1251 AACAAACAACA ATTAGAAAAG AAATTACCTG TAACGTTGA AGATAAGAAG
1301 CGGGAGAACT TTGAACGTGG CAACCTGGAA CTGGAGAAAC GAAGGCAAGC
1351 TCTCCTGGAA CAGCAGCGCA AGGAGCAGGA GCCCCTGGCC CAGCTGGAGC
1401 GGGCGGAGCA GGAGAGGAAG GAGCGTGAGC GCCAGGAGCA AGAGCGCAA
1451 AGACAACCTGG AACTGGAGAA GCAACTGGAA AAGCAGCGGG AGCTAGAACG
1501 GCAGAGAGAG GAGGAGAGGA GGAAAGAAAT TGAGAGGCAGA GAGGCTGCAA
1551 AACGGGAACCT TGAAAGGCAA CGACAACCTG AGTGGGAACG GAATCGAAGG
1601 CAAGAACTAC TAAATCAAAG AAACAAAGAA CAAGAGGACA TAGTTGTACT
1651 GAAAGCAAAG AAAAAGACTT TGGAAATTGA ATTAGAAGCT CTAATGATA
1701 AAAAGCATCA ACTAGAAGGG AAACCTCAAG ATATCAGATG TCGATTGACC
1751 ACCCAAAGGC AAGAAATTGA GAGCACAAAC AAATCTAGAG AGTTGAGAAT
1801 TGCCGAAATC ACCCATCTAC AGCAACAATT ACAGGAATCT CAGCAAATGC
1851 TTGGAAGACT TATTCCAGAA AAACAGATAC TCAATGACCA ATTAAAACAA
1901 GTTCAGCAGA ACAGTTGCA CAGAGATTCA CTTGTTACAC TTAAAAGAGC
1951 CTTAGAAGCA AAAGAACTAG CTCGGCAGCA CCTACGAGAC CAACTGGATG
2001 AAGTGGAGAA AGAAACTAGA TCAAAACTAC AGGAGATTGA TATTTCAAT
2051 AATCAGCTGA AGGAACCTAAG AGAAATACAC AATAAGCAAC AACTCCAGAA

Figure 10

2101 GCAAAAGTCC ATGGAGGCTG AACGACTGAA ACAGAAAGAA CAAGAACGAA
2151 AGATCATAGA ATTAGAAAAA CAAAAGAAG AAGCCAAAG ACGAGCTCAG
2201 GAAAGGGACA AGCAGTGGCT GGAGCATGTG CAGCAGGAGG ACGAGCATCA
2251 GAGACCAAGA AAACCTCCACG AAGAGGAAAA ACTGAAAAGG GAGGAGAGTG
2301 TCAAAAAGAA GGATGGCGAG GAAAAAGGC AACAGGAAGC ACAAGACAAG
2351 CTGGGTCGGC TTTTCCATCA ACACCAAGAA CCAGCTAACG CAGCTGTCCA
2401 GGCACCCCTGG TCCACTGCAG AAAAAGGTCC ACTTACCATT TCTGCACAGG
2451 AAAATGTAAA AGTGGTGTAT TACCGGGCAC TGTACCCCTT TGAATCCAGA
2501 AGCCATGATG AAATCACTAT CCAGCCAGGA GACATAGTCA TGTTGGATGA
2551 AAGCCAAACT GGAGAACCCG GCTGGCTTGG AGGAGAATTA AAAGGAAAGA
2601 CAGGGTGGTT CCCTGCAAAC TATGCAGAGA AAATCCCAGA AAATGAGGTT
2651 CCCGCTCCAG TGAAACCAGT GACTGATTCA ACATCTGCC CTGCCCCCAA
2701 ACTGGCCTTG CGTGAGACCC CGGCCCTTGG GCAGTAACC TCTTCAGAGC
2751 CCTCCACGAC CCCTAATAAC TGGGCCGACT TCAGCTCAC GTGGCCCACC
2801 AGCACGAATG AGAAACCAGA AACGGATAAC TGGGATGCAT GGGCAGCCCA
2851 GCCCTCTCTC ACCGTTCCAA GTGCCGGCCA GTTAAGGCAG AGGTCCGCCT
2901 TTACTCCAGC CACGGCCACT GGCTCCTCCC CGTCTCCTGT GCTAGGCCAG
2951 GGTGAAAAGG TGGAGGGGCT ACAAGCTCAA GCCCTATATC CTTGGAGAGC
3001 CAAAAAAAGAC AACCACTTAA ATTTAACAA AAATGATGTC ATCACCGTCC
3051 TGGAACAGCA AGACATGTGG TGGTTGGAG AAGTTCAAGG TCAGAAGGGT
3101 TGGTCCCCA AGTCTTACGT GAAACTCATT TCAGGGCCCA TAAGGAAGTC
3151 TACAAGCATG GATTCTGGTT CTTCAGAGAG TCCTGCTAGT CTAAAGCGAG
3201 TAGCCTCTCC AGCAGCCAAG CCGGTCGTTT CGGGAGAAGA ATTTATTGCC
3251 ATGTACACTT ACGAGAGTTC TGAGCAAGGA GATTAACTT TTCAGCAAGG
3301 GGATGTGATT TTGGTTACCA AGAAAGATGG TGACTGGTGG ACAGGAACAG
3351 TGGCGACAA GGCGGAGTC TTCCCTCTA ACTATGTGAG GCTTAAAGAT
3401 TCAGAGGGCT CTGGAACTGC TGGGAAAACA GGGAGTTAG GAAAAAAACC
3451 TGAAATTGCC CAGGTTATTG CCTCATACAC CGCCACCGGC CCCGAGCAGC
3501 TCACTCTCGC CCCTGGTCAG CTGATTTGA TCCGAAAAAA GAACCCAGGT
3551 GGATGGTGGG AAGGAGAGCT GCAAGCACGT GGGAAAAGC GCCAGATAGG
3601 CTGGTTCCCA GCTAATTATG TAAAGCTTCT AAGCCCTGGG ACGAGCAAAA
3651 TCACTCCAAC AGAGCCACCT AAGTCAACAG CATTAGCGGC AGTGTGCCAG
3701 GTGATTGGGA TGTACGACTA CACCGCGCAG AATGACGATG AGCTGGCCTT
3751 CAACAAGGGC CAGATCATCA ACGTCTCAA CAAGGAGGAC CCTGACTGGT
3801 GGAAAGGAGA AGTCAATGGA CAAGTGGGGC TCTTCCCATC CAATTATGTG
3851 AAGCTGACCA CAGACATGGA CCCAAGCCAG CAATGAATCA TATGTTGTCC
3901 ATCCCCCCCCT CAGGCTTGAA AGTCCTTTG TGGCTTCCT AGTTACTCAA
3951 ATTGACTTTC CCCCCACCTTT GCACAGGTGC TTCAATAGT TTAAATTTA
4001 TTTTAAATA TATATTAG TTGTTAATA AACAAAATAA ATAATGACT
4051 TCTTGCTAT TTGGTTTTG CAAAAAGACC CACTATCAAG GAATGCTGCA
4101 TGTGCTATTA AAAATTGTTCAAAATGTCCA TAAATCTGAG ACTTGATGTA
4151 TTTTTCTTATTGTTGTCAGTG TTACCAACTA AATTGTGCAG TTTGGGGCTT
4201 TTCCCCCTTA CCATAGAAGT GCAGAGGAGT TCAGTATCTC TGTTTAAAG

Figure 10

09/720934

WO 99/53062

PCT/US99/08371

20 / 30

4251 ACGTATAGAA TGAGCCCAAT TAAAGCGAAG GTGTTGTGC TTGTTGTGT
4301 GTATCAGCTG TACCTTGTG AGCATGTAAT ACATCCTGTA CATAAGAAAT
4351 TAGTTCTTTC CATGGCAAAG CTATTACCTT GTACGATGCT CTAATCATAT
4401 TGCAATTAAAT TTTATTTCGC ACAGTGACCT TGTAGGCCACA TGAGAAAAGCA
4451 CTCTGTGTTT TTGTTCGGTC TCAGATTAT CTGGTTGAGT TGGTGTGTTG
4501 TTTGGGGTTT TTAATTTCGC GTGTTGAT AGCATAAAAT CAGTAGACAA
4551 CACCACTGAG GTCGTTACGA TCAACGATAT CCACAGTCTC TTTTAGTCT
4601 CTGTTACATG AAGTTTATT CCAGTTACTT TTCATGGAAT GACCTATTIT
4651 GAACAAGTAA TTTCTTGAC AAGAAAGAAT GTATAGAAGT CTCCCTGCAA
4701 TTAATTCCA ATGTTTACAT TTTTAACTA GACTGTGGAA TTTCTACAGA
4751 TTAATATGAA ATGGAGCTCA TGGTCCGTT GTGTGTTAGA TATGCTGTAG
4801 CTGAAGCCCT GTTGTCTTT TAAACACTAG TTGGAAGCTC TCAATAAAAA
4851 TGCCTGCTGC TCACAGCACA GAAAATGGGG CAGGGGGAGC CTCAAGCACA
4901 ATCTAGCTGT CCTCCTAAAG ACTCTGTAAT GCTCACTCCC CTCGCCTCT
4951 CCCGGCGCTG TCGGGAGGCT GTGCTGGTGG TCGTGTAGAG GTCCTTCTCC
5001 TTTCACATGG TGCAGAGAGC GAGGACCTCT CCTCCTCGTT CAGTTGCACT
5051 TCAGTATTTT CACGGATATG AATGTAAGAAT ATATAAATAT ATAAACCTGC
5101 GGCTTTAACCA ACTGTAATAC AACCTTTGA ATTAGTTCCG TGTATAGATA
5151 ATTAAATTCT TCATACAAAAA GTTAAAAAAA AAAAAAAAAA AAAAAA

Figure 10

09/720934

PCT/US99/08371

WO 99/53062

21 / 30

Translated Protein Sequence #11

1 MAQFPTPFCCG SLIDIWAIITVE ERAKHDQQFH SLKPISGFIT GDQARNFFFQ
 51 SGLPQPVLAQ IWALADMNND GRMDQVEFSI AMKLKIKLKLQ GYQLPSALPP
 101 VMKQQPV AIS SAPAFGMGGI ASMPPLTAVA PVPMSIPVV GMSPTLVSSV
 151 PTAAVPPLAN GAPPVIQPLP AFAHPAATLP KSSSFSRSGI GSQNLNTKLQK
 201 AQSFDVASVP PVAEWAVPQS SRLKYRQLFN SHDKTMMSGH TGQARTILM
 251 QSSLPQAQLA SIWNLSDIDQ DGKLTAEFFI LAMHLIDVAM SGQPLPPVLP
 301 PEYIPIPPSFR VRSGSICISVI SSTSVQRLP EEPVLEDEQQ QLEKKLPVTF
 351 EDKKRENFER GNLELKRRQ ALLEQQRKEQ ERLAQLERAE QERKERERQE
 401 QERKRQLELE KQLEKQRELE RQREERRKE IERREAAKRE LERQRQLEWE
 451 RNRRRQELLNQ RNKEQEDIVV LAKAKKKTLEF ELEALNDKKH QLEGKLQDIR
 501 CRLTTORQEI ESTNKSREL R IAEITHLQQQ LQESQQMILGR LIPEKQILND
 551 QLKQVQQNSL HRDLSVTLK ALEAKELARO HLRDOLDEVE KETRSKLQEI
 601 DIFNNQLKEL REIHNKQQLQ KQKSMEAERL KQKEQERKII ELEKQKEEAQ
 651 RRAQERDKQW LEHVQQEDEH QRPRKLHEEE KLKREESVKK KDGEEKGKQE
 701 AQDKLGLRFH QHQEPAKPAV QAPWSTAEGK PLTISAQENV KVYYYRALYP
 751 FESRSHDEIT IQPGDIVMVD ESQTGEPEGLV GGEALKGKTGW FPANYAEKIP
 801 ENEVPAPVKP VTDSTSAPAP KLALRETPAP LAVTSSSEPST TPPNWADFSS
 851 TWPTSTNEKP ETDNWDAWAA QSLTVPSAG QLRQRSAFTP ATATGSSPSP
 901 VLGQGEKVEG LQAQALYPWR AKKDNLHNFN KNDVITVLEQ QDMWWFGEVQ
 951 GQKGWFPKSY VKLISGPIRK STMSDGSSE SPAASKRVAS PAAKPVVSGE
 1001 EFIA MYTYES SEQGDLTFQQ GDVILVTKKD GDWWTGTVGD KAGVFPNSNYV
 1051 RLKDSESGGT AGKTGSLGKX PEIAQVIASY TATGPEQLTL APGQLLIRK
 1101 KNPGGWWEGE LQARGKKRQI GWFPANYVKL LSPGTSKITP TEPPKSTALA
 1151 AVCQVIGMYD YTAQNDDELA FNKGQIINV L NKEDPDWWKG EVNGQVGLFP
 1201 SNYVKLTTDM DPSQQ*

whole protein sequence:

1 EWRRQGRERS LVAP*YGGSR GRIPSGLRDG QRGGRGWCAG LRLRLPSQRR
 51 VSGTDLSLGR QRGPARR*GV D*QGKSNRTM AQFPPTFGGS LDIWAITVEE
 101 RAKHDQQFHS LKPIISGFITG DQARNFFFQ GLPQPVLAQI WALADMNNDG
 151 RMDQVEFSIA MKL:IKLKLQG YQLPSALPP MKQQPVAISS APAFGMGGIA
 201 SMPPLTAVAP VPMGSIPTVG MSPTLVSSV TAAVPPPLANG APPVIQPLPA
 251 FAHPAATLPK SSSFSRSGPG SQLNTKLQKA QSFDVASVPP VAEWAVPQSS
 301 RLKYRQLFNS HDKTMMSGHT GPOARTILMQ SSSLPAQLAS IWNLSIDIDQD
 351 GKLTAEFFI AMHLIDVAM GQPLPPVLP EYIPPSFRRV RSGSGCIVIS
 401 STSVDQRLPE EPVLEDEQQQ LEKKLPVTE DKKRENFERG NLELEKRRQA
 451 LLEQQRKEQE RLAQLERAEQ ERKERERQE ERKRQLELEK QLEKQRELER
 501 QREEERRKEI ERREAAKREL EROROLEWER NRQEELLNQR NKEQEDIVVL
 551 KAKKKTLEFE LEALNDKKHQ LECKLQDIRC RLTTRQREIE STNKSRELRI
 601 AEITHLQQQL QESQQMLGRL IPEKOILNDQ LKQVQQNSLH RDSLVTLKRA
 651 LEAKELARQH LRDQLDEVEK ETRSKLQEI IFNNQLKELR EIHNKQQLQK
 701 QKSMEAERLK QKEQERKII LEKOKEEEAQR RAQERDKQWL EHVVQQEDEHQ
 751 RPRKLHEEEK LKREESVKK DGEEKGKQEA QDKLGLRFHQ HQEPAKPAVQ
 801 APWSTAEGK CP LTISAQENVK VVYYYRALYPF ERSRSHDEITI QPGDIVMVE
 851 SQTGEPEGLV GELKGKTGW PANYAEKIP NEVPAPVKP TDSTSAPAPK
 901 LALRETPAPL AVTSSSEPST TPPNWADFSST WPTSTNEKP ETDNWDAWAAQ
 951 PSLTVPSSAQG LRQRSAFTP TATGSSSPV LGQGEKVEGL QAAQALYPWRA
 1001 KKDNHLMFNK NDVITVLEQQ DMWVFGEVQ QKGWFPKSYV KLIISGPIRK
 1051 TSMDSGSSSES PASLKRVASP AAQPVVSGEE FLAMYTYESS EQGDLTFQQG
 1101 DVILVTKKD DWWTGTVGDK AGVFPNSNYV LKDSESGGT A GKTGSLGKKP
 1151 EIAQVIASYT ATGPEQLTLA PGQLLIRKX NPGWWEGEL QARGKKRQIG
 1201 WFPANYVKKL SPGTSKITP EPPKSTALAA VCQVIGMYD TAQNDDELAF
 1251 NKGQIINV L NKEDPDWWKG EVNGQVGLFPs NYVKLTTDM PSQQ*ICCP
 1301 SPPQA*KSF C GFPYSN*LS PTFAQVLSIV LKLFNLYIFS FLINKINKL
 1351 LCYFGFAKRP TIKECCMCY* KLFOMSINLR LDVFFFHFVQC YOLNCAVWGF
 1401 SPLP*KCRGV QYLCFKD*V N EPN*SEGVC CLCVSAVPC* ACNTSCT*EI
 1451 SSFHGKAITL YDALIILHLL LFCTVTL*PH EKALCVFVR S QIYLVELVFC
 1501 LGFLILRVCIA *NQ*TTPLR SLRSTISTVS F*SLLHEVLF QLLFME*PIL
 1551 NK*FS*QERM YRSLPAINFQ CLHFLTRLWN FYRLI*NGAH GPFVC*ICCS
 1601 *SPVCLLNNTS WKLSIKMPAA HSTENGAGGA SSTI*LSS*R LCNAHSPRVL
 1651 PALSGGCAGG RVEVLLLSHG AESEDLSSSF SCTSVFSRI* M*N*IYKPA
 1701 ALTTVIQFPE LVPCION*IL HTKVKKKKKK K

Figure 11

09/720934

PCT/US99/08371

WO 99/53062

22 / 30

1 CGGGGATGGT GTGCCGGGCT GCGGCTCCTG CGTCCCTCCC AGCGGCGCGT
51 GAGCGGCACT GATTGCCCC TGCGGCGGCA GCGCGGACCC GCGCGGAGAT
101 GAGGCCTCGA TTAGCAAGGT AAAAGTAACA GAACCATGGC TCAGTTCCA
151 ACACCTTTG GTGGCAGCCT GGATATCTGG GCCATAACTG TAGAGGAAG
201 AGCGAACAT GATCAGCAGT TCCATAGTTT AAAGCCAATA TCTGGATTCA
251 TTACTGGTGA TCAAGCTAGA AACTTTTTT TTCAATCTGG GTTACCTCAA
301 CCTGTTTAG CACAGATATG GGCAGTAGCT GACATGAATA ATGATGGAAG
351 AATGGATCAA GTGGAGTTT CCATAGCTAT GAAACTTATC AACTGAAGC
401 TACAAGGATA TCAGCTACCC TCTGCACTTC CCCCTGTCA GAAACAGCAA
451 CCAGTTGCTA TTTCTAGCGC ACCAGCATT GGTATGGGAG GTATGCCAG
501 CATGCCACCG CTTACAGCTG TTGCTCCAGT GCCAATGGGA TCCATTCCAG
551 TTGTTGGAAT GTCTCCAACC CTAGTATCTT CTGTTCCAC AGCAGCTGTG
601 CCCCCCCTGG CTAACGGGGC TCCCCCTGTT ATACAACCTC TGCCTGCATT
651 TGCTCATCCT GCAGCCACAT TGCCAAAGAG TTCTTCTTT AGTAGATCTG
701 GTCCAGGGTC ACAACTAAC ACTAAATTAC AAAAGGCACA GTCATTGAT
751 GTGGCCAGTG TCCCACCAAGT GGCAGAGTGG GCTGTTCTC AGTCATCAAG
801 ACTGAAATAC AGGCAATTAT TCAATAGTCA TGACAAAAT ATGAGTGGAC
851 ACTTAACAGG TCCCAAGCA AGAACTATTG TTATGCAGTC AAGTTACCA
901 CAGGCTCAGC TGGCTCAAT ATGGAATCTT TCTGACATTG ATCAAGATGG
951 AAAACTTACA GCAGAGGAAT TTATCCTGGC AATGCACCTC ATTGATGTAG
1001 CTATGTCTGG CCAACCACTG CCACCTGTCC TGCCCTCCAGA ATACATTCCA
1051 CCTTCTTTA GAAGAGTTG ATCTGGCACT GGTATATCTG TCATAAGCTC
1101 AACATCTGTA GATCAGAGGC TACCAGAGGA ACCAGTTTA GAAGATGAAC
1151 AACAAACAATT AGAAAAGAAA TTACCTGTA CGTTGAAGA TAAGAAGCGG
1201 GAGAACTTTG AACGTGGCAA CCTGGAACCTG GAGAAACGAA GGCAAGCTCT
1251 CCTGGAACAG CAGCGCAAGG AGCAGGAGCG CCTGGCCCAAG CTGGAGCGGG
1301 CGGAGCAGGA GAGGAAGGAG CGTGAGCGCC AGGAGCAAGA GCGAAAAGA
1351 CAACTGGAAC TGGAGAAGCA ACTGGAAAAG CAGCGGGAGC TAGAACGGCA
1401 GAGAGAGGAG GAGAGGAGGA AAGAAAATTGA GAGGCAGAGAG GCTGCAAAAC
1451 GGGAACTTGA AAGGCAACGA CAACTTGAGT GGGAACGGAA TCGAAGGCAA
1501 GAACTACTAA ATCAAAGAAA CAAAGAACAA GAGGACATAG TTGTACTGAA
1551 AGCAAAGAAA AAGACTTTGG AATTGAAATT AGAAGCTCTA AATGATAAAA
1601 AGCATCAACT AGAAGGGAAA CTTCAAGATA TCAGATGTG ATTGACCACC
1651 CAAAGGCAAG AAATTGAGAG CACAAACAAA TCTAGAGAGT TGAGAATTGC
1701 CGAAATCACCC CATCTACAGC AACAAATTACA GGAATCTCAG CAAATGCTTG
1751 GAAGACTTAT TCCAGAAAAA CAGATACTCA ATGACCAATT AAAACAAGTT
1801 CAGCAGAACCA GTTGCACAG AGATTCACTT GTTACACTTA AAAGAGCCTT
1851 AGAAGCAAAA GAACTAGCTC GGCAGCACCT ACGAGACCAA CTGGATGAAG
1901 TGGAGAAAGA AACTAGATCA AAACATACAGG AGATTGATAT TTCAATAAT
1951 CAGCTGAAGG AACTAAGAGA AATACACAAAT AAGCAACAAAC TCCAGAAGCA
2001 AAAGTCCATG GAGGCTGAAC GACTGAAACA GAAAGAACAA GAACGAAAGA
2051 TCATAGAATT AGAAAAAAAAA AAAAAAAAAA

Figure 12

09/72093

#5 translated Protein sequence:

1 MAQFPTPGG SLDIWAITVE ERAKHDQQFH SLKPISGFIT GDQARNFFFQ
51 SGLPQPVLAQ IWALADMNND GRMDQVEFSI AMKLIKLLQ GYQLPSALPP
101 VMKQQPVAIS SAPAFGMGGI ASMPLTAVA PVPMSGIPVV GMSPTLVSSV
151 PTAAVPLAN GAPPVIQPLP AFAHPAATLP KSSSFSRSGP GSQNLNTKLQK
201 AQSFDVASVP PVAEWAVPQS SRLKYRQLFN SHDKTMSGHL TGPQARTILM
251 QSSLMPQAQLA SIWNLSDIDQ DGKLTAEEFI LAMHLLIDVAM SGQPLPPVLP
301 PEYIPPSFRR VRSGSGISVI SSTSVVDQRLP EEPVLEDEQQ QLEKKLPVTF
351 EDKKRENFER GNLELEKRRQ ALLEQQRKEQ ERLAQLERAQ QERKERERQE
401 QERKRQLELE KQLEKQRELE RQREEERRKE IERREAQKRE LERQRQLEWE
451 RNRRQELLNQ RNKEQEDIVV LKAKKKTLEF ELEALNDKKH QLEGKLQDIR
501 CRLTTQRQEI ESTNKSRELR IAEITHLQQQ LQESQQMLGR LIPEKQILND
551 QLKQVQQNSL HRDSLVTLKR ALEAKELARQ HLRDQLDEVE KETRSKLQEI
601 DIFNNQLKEL REIHNKQQLQ KQKSMEAERL KQKEQERKII ELEKKKKK

whole sequence

1 RGWCAGLRLL RPSQRVS GT DLSLGRQRGP ARR*GVD*QG KSNRTMAQFP
51 TPFGGSLDIW AITVEERAHK DQQFHSLKPI SGFITGDQAR NFFFQSGLPQ
101 PVLAQIWALA DMNNNDGRMDQ VEFSIAMKLI KLKLQGYQLP SALPPVMKQQ
151 PVAISSAPAF GMGGIASMPP LTAVAPVPMG SIPVVGMSPT LVSSVPTAAV
201 PPLANGAPPV IQPLPAFAHP AATLPKSSSF SRSGPGSQLN TKLQKAQSF
251 VASVPPVAEW AVPQSSRLKY RQLFNSHDKT MSGHILTGPQA RTILMQSSL
301 QAQLASIWNL SDIDQDGKLT AEEFILAMHL IDVAMSGQPL PPVLPPYEYIP
351 PSFRRVRSGS GISVISSTSV DQR LPPEPV EDEQQQLEKK LPVTFEDKKR
401 ENFERGNLEL EKRRQALLEQ QRKEQERLAQ LERAQERKE RERQEQERKR
451 QLELEKQLEK QRELERQREE ERRKEIERRE AAKRELERQR QLEWERNRRQ
501 ELLNQRNKEQ EDIVVLKAKK KTLEFELEAL NDKKHQLEGK LQDIRCRLTT
551 QRQEIESTNK SRELRIAETI HLQQQLQESQ QMLGRLIPEK QILNDQLKQV
601 QQNSLHRDSL VTLKRALEAK ELARQHLDQ LDEVEKETRS KLQEIDIFNN
651 QLKELREIHN KQQLQKQKSM EAERLKQKEQ ERKIELEKK KKK

Figure 13

09/720934

WO 99/53062

PCT/US99/08371

24 / 30

1 GACCACCCAA AGGCAAGAAA TTGAGAGCAC AAACAAATCT AGAGAGTTGA
51 GAATTGCCGA AATCACCCAT CTACAGCAAC ATTACAGGA ATCTCAGCAA
101 ATGCTTGGAA GACTTATTCC AGAAAAACAG ATACTCAATG ACCAATTAAA
151 ACAAGTTCAAG CAGAACAGTT TGCACAGAGA TTCACTTGTGTT ACACCTAAAA
201 GAGCCTTAGA AGCAAAAGAA CTAGCTCGGC AGCACCTACG AGACCAACTG
251 GATGAAGTGG AGAAAGAAC TAGATCAAAA CTACAGGAGA TTGATATTT
301 CAATAATCAG CTGAAGGAAC TAAGAGAAAT ACACAATAAG CAACAACCTCC
351 AGAACAAAAA GTCCATGGAG GCTGAACGAC TGAAACAGAA AGAACAAAGAA
401 CGAAAGATCA TAGAATTAGA AAAACAAAAA GAAGAACGCCC AAAGACGAGC
451 TCAGGAAAGG GACAAGCAGT GGCTGGAGCA TGTGCAGCAG GAGGACGAGC
501 ATCAGAGACC AAGAAAACTC CACGAAGAGG AAAAATGAA AAGGGAGGAG
551 AGTGTAAAAA AGAAGGATGG CGAGGAAAAA GGCAAAACAGG AAGCACAAGA
601 CAAGCTGGGT CGGCTTTCC ATCAACACCA AGAACAGCT AAGCCAGCTG
651 TCCAGGCACC CTGGTCCACT GCAGAAAAAG GTCCACTTAC CATTCTGCA
701 CAGGAAAATG TAAAAGTGGT GTATTACCGG GCACTGTACC CCTTTGAATC
751 CAGAACCAT GATGAAATCA CTATCCAGCC AGGAGACATA GTCATGGTGG
801 ATGAAAGCCA AACTGGAGAA CCCGGCTGGC TTGGAGGAGA ATAAAAGGA
851 AAGACAGGGT GGTTCCCTGC AAACATATGCA GAGAAAATCC CAGAAAATGA
901 GGTTCCCGCT CCAGTGAAAC CAGTGACTGA TTCAACATCT GCCCCTGCC
951 CCAAACCTGGC CTTGCGTGAG ACCCCCCGCC CTTTGGCAGT AACCTCTCA
1001 GAGCCCTCCA CGACCCCTAA TAACTGGGCC GACTTCAGCT CCACGTGGCC
1051 CACCAGCAGC AATGAGAAC CAGAACCGGA TAACTGGGAT GCATGGCAG
1101 CCCAGCCCTC TCTCACCGTT CCAAGTGGCG GCCAGTTAAG GCAGAGGTCC
1151 GCCTTACTC CAGCCACGGC CACTGGCTCC TCCCCGTCTC CTGTGCTAGG
1201 CCAGGGTGAA AAGGTGGAGG GGCTACAGC TCAAGCCCTA TATCCTTGG
1251 GAGCCAAAAAA AGACAACCAC TAAATTITA ACAAAAATGA TGTCACTCACC
1301 GTCCTGGAAC AGCAAGACAT GTGGTGGTTT GGAGAACGTT AAGGTAGAA
1351 GGGTTGGTTC CCCAAGTCTT ACGTGAACACT CATTCAAGGG CCCATAAGGA
1401 AGTCTACAAG CATGGATTCT GTTCTTCAG AGAGTCCTGC TAGTCTAAAG
1451 CGAGTAGCCT CTCCAGCAGC CAAGCCGGTC GTTCCGGAG AAGAAATTGC
1501 CCAGGTTATT GCCTCATACA CCGCCACCCGG CCCCAGCAG CTCACTCTCG
1551 CCCCTGGTCA GCTGATTTG ATCCGAAAAAA AGAACCCAGG TGGATGGTGG
1601 GAAGGAGAGC TGCAAGCAGC TGGGAAAAAG CGCCAGATAG GCTGGTCCC
1651 AGCTAATTAT GTAAAGCTTC TAAGCCCTGG GACGAGCAAA ATCACTCCAA
1701 CAGAGCCACC TAAGTCACCA GCATTAGCGG CAGTGTGCCA GGTGATTGGG
1751 ATGTACGACT ACACCGCGCA GAATGACGAT GAGCTGGCCT TCAACAAGGG
1801 CCAGATCATC AACGTCTCA ACAAGGAGGA CCTGACTGG TGGAAAGGAG
1851 AAGTCAATGG ACAAGTGGGG CTCTTCCCAT CCAATTATGT GAAGCTGACC
1901 ACAGACATGG ACCCAAGCCA GCAATGAATC ATATGTTGTC CATCCCCCCC
1951 TCAGGGCTTGA AAGTCCTTT GTGGCTTCC TAGTTACTCA AATTGACTTT
2001 CCCCCACCTT TGCACAGGTG CTTCAATAG TTTAAAATT ATTTTAAAT

Figure 14

09/720934

PCT/US99/08371

WO 99/53062

25 / 30

2051 ATATATTTA GCTTTTAAT AAACAAAATA AATAAATGAC TTCTTGCTA
2101 TTTGGTTT GCAAAAGAC CCACTATCAA GGAATGCTGC ATGTGCTATT
2151 AAAAATTGTT CCAAATGTCC ATAAATCTGA GACTTGATGT ATTTTTCAT
2201 TTTGTCCAGT GTTACCAACT AAATTGTGCA GTTGGGGCT TTTCCCCCTT
2251 ACCATAGAAG TGCAGAGGAG TTCAGTATCT CTGTTTAAA GACGTATAGA
2301 ATGAGCCAA TTAAAGCGAA GGTGTTGTG CTTGTTGTG TGTATCAGCT
2351 GTACCTTGTT GAGCATGTA TACATCCTGT ACATAAGAAA TTAGTTCTT
2401 CCATGGCAAA GCTATTACCT TGTACGATGC TCTAATCATA TTGCATTAA
2451 TTTTATTTG CACAGTGACC TTGAGCCAC ATGAGAAAGC ACTCTGTGTT
2501 TTTGTCGGT CTCAGATTAA TCTGGTTGAG TTGGTGTGTTT GTTTGGGGTT
2551 TTTAATTTG CGTGTGCA TAGCATAAAA TCAGTAGACA ACACCACTGA
2601 GGTGTTACG ATCAACGATA TCCACAGTCT CTTTTAGTC TCTGTTACAT
2651 GAAGTTTAT TCCAGTTACT TTTCATGGAA TGACCTATT TGAACAAGTA
2701 ATTTTCTTGA CAAGAAAGAA TGTATAGAAG TCTCCCTGCA ATTAATTCC
2751 AATGTTACA TTTTTAACT AGACTGTGGA ATTCTACAG ATTAATATGA
2801 AATGGAGCTC ATGGTCCGTT TGTGTGTTAG ATATGCTGTA GCTGAAGCCC
2851 TGTTTGTCTT TTAAACACTA GTTGGAAAGCT CTCATAAAA ATGCCTGCTG
2901 CTCACAGCAC AGAAAATGGG GCAGGGGGAG CCTCAAGCAC AATCTAGCTG
2951 TCCTCCTAAA GACTCTGTAA TGCTCACTCC CCTCGCGTTC TCCC GGCGCT
3001 GTCGGGAGGC TGTGCTGGTG GTCGTGTAAG GTCCCTCTCC TTTCACATGG
3051 TGCAGAGAGC GAGGACCTCT CCTCCTCGTT CAGTTGCACT TCAGTATTT
3101 CACGGATATG AATGTAAAAT ATATAAATAT ATAAACCTGC GGCTTTAACAA
3151 ACTGTAATAC AACTTTGA ATTAGITCCG TGTATAGATA ATTAAATTCT
3201 TCATACAAAAA GTTAAAAAAA AAAAAAAAAA A

Figure 14

#9 translated protein sequence:

1 TTQRQEIEST NKSRELRIA E ITHLQQQLQE SQQMLGRLIP EKQILNDQLK
 51 QVQQNSLHRD SLVTLKRALE AKELARQHLR DQLDEVEKET RSKLQEIDIF
 101 NNQLKELREI HNKQQLQKQK SMEAERLKQK EGERKIIIELE KQKEEAQRRA
 151 QERDKQWLEH VQQEDEHQRP RKLHEEEKLK REESVKKKD E EKKGKQEAQD
 201 KLGRLFHQHQ EPAKPAVQAP WSTAEGPLT ISAQENVKV Y YRALYPFES
 251 RSHDEITIQP GDIVMVDESQ TGEPEGWLGGE LKGKTGWFP NYAEKIPENE
 301 VPAPVKPVTD STSAPAPKLA LRETPAPALAV TSSEPSTTPN NWADFSSTWP
 351 TSTNEKPETD NWDAWAQPS LTVPSAGQLR QRSAFTPATA TGSSPSPVLG
 401 QGEKVEGLQA QALYPWRACK DHNLNFNKND VITVLEQQDM WWFGEVQGQK
 451 GWFPKSYVKL ISGPIRKSTS MDSGSSES P A SLKRVASPAA KPVVS GEEIA
 501 QVIASYTATG PEQLTLAPGQ LILIRKKNP G GWWE GELQAR GKKRQIGWFP
 551 ANYVKLLSPG TSKITPTTEPP KSTALAAVCQ VIGMYDYTAQ NDDELAFNKG
 601 QIINVLNKED PDWWKGEVNG QVGLFPSNYV KLTTDMDP SQ Q*

Whole protein sequence

1 TTQRQEIEST NKSRELRIA E ITHLQQQLQE SQQMLGRLIP EKQILNDQLK
 51 QVQQNSLHRD SLVTLKRALE AKELARQHLR DQLDEVEKET RSKLQEIDIF
 101 NNQLKELREI HNKQQLQKQK SMEAERLKQK EGERKIIIELE KQKEEAQRRA
 151 QERDKQWLEH VQQEDEHQRP RKLHEEEKLK REESVKKKD E EKKGKQEAQD
 201 KLGRLFHQHQ EPAKPAVQAP WSTAEGPLT ISAQENVKV Y YRALYPFES
 251 RSHDEITIQP GDIVMVDESQ TGEPEGWLGGE LKGKTGWFP NYAEKIPENE
 301 VPAPVKPVTD STSAPAPKLA LRETPAPALAV TSSEPSTTPN NWADFSSTWP
 351 TSTNEKPETD NWDAWAQPS LTVPSAGQLR QRSAFTPATA TGSSPSPVLG
 401 QGEKVEGLQA QALYPWRACK DHNLNFNKND VITVLEQQDM WWFGEVQGQK
 451 GWFPKSYVKL ISGPIRKSTS MDSGSSES P A SLKRVASPAA KPVVS GEEIA
 501 QVIASYTATG PEQLTLAPGQ LILIRKKNP G GWWE GELQAR GKKRQIGWFP
 551 ANYVKLLSPG TSKITPTTEPP KSTALAAVCQ VIGMYDYTAQ NDDELAFNKG
 601 QIINVLNKED PDWWKGEVNG QVGLFPSNYV KLTTDMDP SQ Q*IICCPSP
 651 QA*KSFCGFP SYSN*LSPTF AQVLSIVLKL FLNIYFSFLI NKINK*LLCY
 701 FGFAKRPTIK ECCMCY*KLF QMSINLRDV FFHFVQCYQL NCAVWGFSPL
 751 P*KCRGVQYL CFKDV*NEPN *SEGVCACLC VSAVPC*ACN TSCT*EISSF
 801 HGKA ITLYDA LIILHLILFC TVTL*PHEKA LCVFVRSQLY LVELVFCLGF
 851 LILRCVIA*N Q*TTPRLSLR STISTVSF*S LLHEVLFQLL FME*PILNK*
 901 FS*QERMYRS LPAINFQCLH FLTRLWNFYR LI*NGAHGPF VC*ICCS*SP
 951 VCLLN TSWKL SIKMPAAHST ENGAGGASST I*LSS*RLCN AHSPRVLPAL
 1001 SGGCAGGRV SFSFHMVQRA RTSPPRSVAL QYFHGYECKI YKYINLRL*Q
 1051 L*YNLLN*FR V*IICFFIQK LKKKKK

Figure 15



Mouse E9
Tissue

Embryo day 9

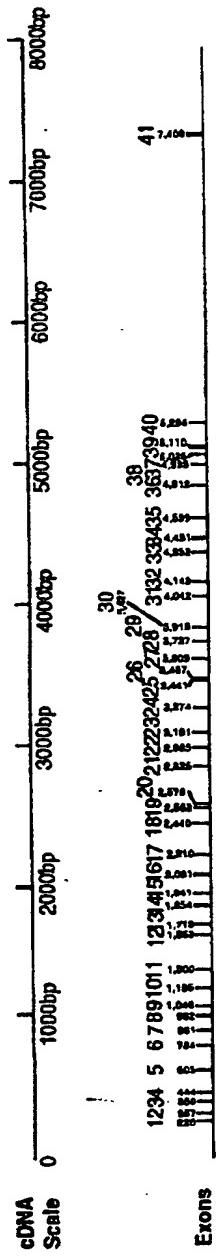
Figure 16

2320-1-001 PCR

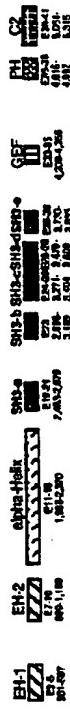
(Sheet 28 of 30)

Summary of Studies on ITS (Intersectin) AKA SH3P17

I. Gene Sequence :



II. Protein Domains vs. Nucleotide sequence:



III. Gene Expression of Human Adult and Fetal Tissues:

Probes used	15.21A 2.0Kb	20.25K 2.0Kb	21.16A 1.6Kb	21.16B 1.6Kb	21.16C 1.6Kb	21.16D 1.6Kb	21.16E 1.6Kb
Northern Blots	-	-	-	-	-	-	-

15Kb (+) - - - - - -
 9.0Kb (+FB) - - - - - -
 5.4Kb + + + + + + +
 4.5Kb (+FL) (+FL) (+FL) (+FL) (+FL) (+FL) (+FL)
 2.0Kb + + + + - - +

*Human ITS (Intersectin), AKA SH3P17 is ubiquitously expressed with extensive alternative splicing generating tissue and developmental stage-specific expression.

IV. Gene Expression with Antibodies to SH3e:

- Gene expression is specific to subpopulation of neurons during CNS morphogenesis and in fetal liver, suggesting possible roles for this gene in hematopoiesis, possibly leukemia and platelet formation as well as in brain formation.

B= band seen only in adult and fetal brain
 AB= band seen only in adult brain
 FB= band seen only in fetal brain
 FL= band seen only in fetal liver

Figure 17

2320-1-001 PCT

(Sheet 29 of 30)

Figure 18

09/72093

PCT/US99/08371

2320-1-001 PCT (Sheet 30 of 30)

LUTZER KREPLINESZ TEREVÍTÉNKÖLÖK MICH! I AL KERÉKRSZKEPHTMIG GÖL SJOG PHÍGYPÍR EREZSÍL KEGYEDLÍSTY KÖRNYECKI PÁLYÁKNALENT PERIODISTAI KÖHL EKKEHC

	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140				
Menor	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140			
Dep 10	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146		
SPD 17	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152
Constitut	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	

Kenop
Dup 1
stop
frozen

Figure 18